MIRADA
The Imaging Software People

The cutting edge: Delineating contours with Deep Learning

P. Aljabar and M.J Gooding

Introduction

Processing and analyzing medical images and clinical data, with the automation provided by statistical or machine learning methods, are well-established components of diagnostic and treatment pathways. Such automation, whether partial or full, assists clinicians and healthcare experts in an ever-widening range of tasks, informing diagnosis, planning and decision-making. Through the introduction of deep learning methods, a more diverse range of applications is being tackled, ranging from the detection and tracking of skin cancer lesions [1] through to diagnostic support for Alzheimer's disease [2] and the prediction of reporting descriptions from imaging data [3]. Deep learning techniques, a subset of machine learning that use neural network models, have undergone a recent surge in usage and interest due to improvements in technology and the ready availability of large datasets. Figure 1 shows *PubMed* data on the numbers of published papers that concern machine learning and deep learning. The growth in the development of deep learning methods in a clinical setting is particularly rapid, giving an indication of the current and future power of these methods to address both clinical needs and research questions.

There are many ways in which deep learning can be used in healthcare and this paper will outline some of the general concepts of deep learning and then look at its application to the specific task of delineation, the generation of pixel-wise segmentation of organs and structures in medical image data.

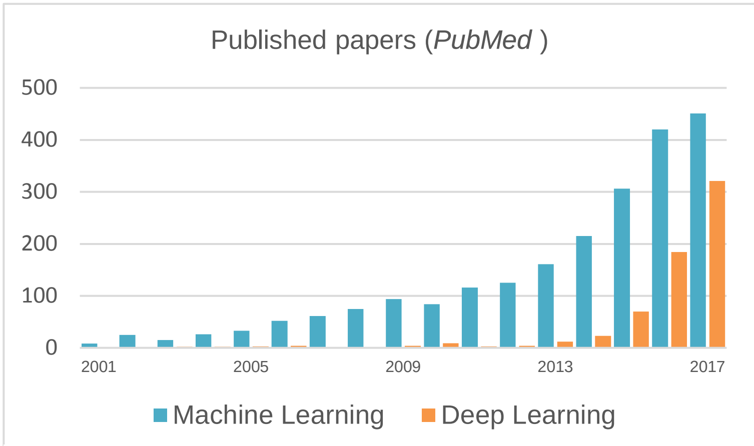

Figure 1: Publication counts on PubMed on Machine Learning and Deep Learning in recent years.

What is deep learning?

A simple example of a learning method is regression, where one tries to predict an output variable y from an input variable x by using a linear model of the form $y = ax+b$. In this case, a and b are the model parameters. A researcher in heart disease, for example, can collect data from a patient group and try to determine whether body mass index (x) can be used to predict cholesterol level (y). Given a set of measurements (x_1,y_1), (x_2,y_2), ... a linear regression model seeks parameters a and b that are optimal in the sense that the differences between predicted and observed measurements are minimized. i.e. given any sample measurement, the predicted value based on BMI, ax_k+b, is as close as possible to the observed cholesterol level y_k. This simple example embodies many machine learning methods where we seek to use input data to predict a particular output, and these methods fall into the general category of supervised learning. The linear regression model is perhaps the simplest example of a large number of possible models that are used in machine learning.

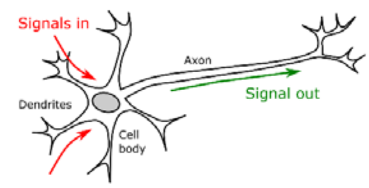

Figure 2: Schematic illustration of a neuron.

One important subset of machine learning models were inspired by the behavior of neurons in the brain, in particular by the way that they can receive signals from other neurons and in turn pass these signals on to other neurons (Fig. 2). Neuronal behavior can be modelled in software and, in its simplest form, an artificial neural network (ANN) consists of software 'neurons' arranged in the following 'layers': a layer to receive the input signal, a layer to deliver the output, and one or more intermediate 'hidden' layers (Fig. 3). Each layer receives signals from the previous layer, processes them, and passes the results to the next layer. A network that has many layers is described as 'deep', which leads to the term 'deep learning', i.e. the process of learning the optimal neuronal parameters to produce the best outputs for the given inputs to the network (in the same way as for the linear regression above).

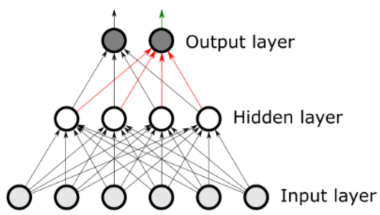

Figure 3: Architecture of a simple artificial neural network.

How can deep learning be applied to organ-at-risk contouring?

Many neuron models have been developed with various systems, known as architectures, by which they are collected into layers and interconnected. Neural network architectures have been adapted for use with diverse types of input data, from share prices in stock markets to audio streams of natural speech. Image processing tasks have been a particular focus of active research into neural networks and convolutional neural networks (CNNs) in particular emerged as a powerful tool for learning from image data. Convolution is a standard image processing technique for generating feature maps from images, such as an edge map. This is achieved by calculating the 'response' values of a set of weights in a kernel as it slides across an image (Fig. 4).

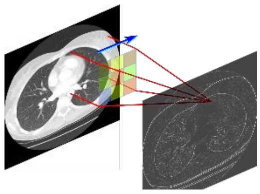

Figure 4: Convolution of an image to generate a feature map

Convolution's incorporation into deep learning frameworks yielded very impressive results. A key aspect of their power lies in their ability to learn image features at multiple scales that are directly relevant to the learning task in hand – this avoids the need for the user to try and identify good features prospectively.

The first notable success for CNNs came in 1998 in handwritten digit recognition (Fig. 5) [4]. Advances in technology and optimization meant that, by 2012, a CNN with very a similar (albeit larger) architecture was able to identify 1000 objects with great accuracy in a one million image database [5].

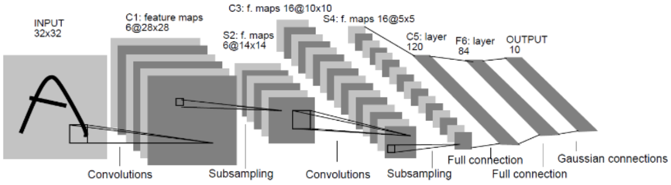

Figure 5: Architecture of a convolution neural network for digit recognition.

Determining a digit or object that an image contains involves assigning a category or label, i.e. it is a classification task where a single label is applied to the whole image. When contouring structures in medical images, we need to assign a separate label to every pixel (or voxel) within the image, for example 'heart', 'liver' or 'background'. The flexibility of CNNs makes them readily adaptable to this task and they have been successfully applied to the segmentation of a variety of structures in different image modalities, such as knee-cartilage [6] and brain structures [7] in MR, organs-at-risk [8] or pulmonary nodules [9] in CT, and cell boundaries in high-resolution microscopy data [10]. At their core, all such convolutional approaches for the dense pixel-wise segmentation of structures in images require training data in the form of images along with ground truth delineations, ideally rated by clinicians with anatomical expertise. Further aspects of the data needed are discussed below.

The advantages of deep learning over current contouring methods

Current commercial solutions for automatically contouring structures in medical images are dominated by atlas-based segmentation and active shape models. Atlas-based segmentation requires a set of images with corresponding ground truth labels (atlases) to be shipped as part of the software. To estimate a contour on new data, the atlases need to be co-registered with the target image to be segmented, and their labels are propagated and combined. One limitation arises when the target image has a very different anatomy from a typical atlas, the registration method may not be able to accurately align atlases to the target, ultimately leading to a loss of segmentation accuracy. Additionally, atlas-based contouring does not scale well with the number of atlases; registering a large number is computationally demanding. In practice, atlas sets are typically kept small (10-20), or schemes are used to pre-select atlases suitable for the target, although an effective selection method is yet to be established [11]. Deep learning models have the capacity and complexity to learn from very large data sets by optimizing a significant number of parameters during training ($\sim 10^6$-10^8). While this training can require significant computation, application of the trained deep neural network to new data is rapid and does not require any data to be shipped with the model – the network's parameters encode how structures should be contoured. Active shape models (ASMs) are also parametric but, in comparison with deep models, have far fewer effective degrees of freedom (due to strong regularization) and do not match the representational power complexity of neural networks. Deep networks also operate directly in the image data and are not restricted by the topology of the of the statistical shape model in an ASM.

Method comparison:	Atlas-based	Active shape models	Deep Learning
Effective degrees of freedom	$\sim 1\times10^5$	$\sim 4\times10^5$	$\sim 1\times10^6$-1×10^8
Topologically restricted to input data?	Yes	Yes	No
Dependent on registration	Yes	Yes	No
Scales with training data (at run-time)	No	Yes	Yes
Requires shipping of data	Yes	No	No
Explicit parameters in model	No	Yes	Yes

What data, and how much, are required?

Deep networks have a great capacity to learn from example segmentations in medical images. With so many parameters, a significant number – hundreds or even thousands of images – can be required to train a deep network. However, this can be reduced to around one hundred, or even fewer, through techniques such as data augmentation (generating transformed versions of observed datasets) or the use of pre-trained networks (with parameters that were previously learned from other data). An important point is that training data should contain the natural anatomical variation expected in the patient population. This leads to improved accuracy when applying the learned model to new data because similar examples will have been 'seen' during training – something that is more challenging for the smaller datasets used in atlas-based methods.

With sufficient data, deep learning methods can closely imitate human segmentation behavior and, when bench-marked against the level of inter-rater variability, can be shown to provide contours with greater levels of clinical acceptance (See the separate Mirada white paper on Thoracic OAR contouring).

Conclusion

Deep learning uses neural network models with very large numbers of parameters that are capable of representing complex and rich functions. They can be used for contouring structures in medical images and, in contrast with previous generations of auto-contouring approaches, can produce contours with a greater level of clinical acceptance, leading to significant time-savings for this part of the radiotherapy treatment planning process.

References

1. Esteva A, Kuprel B, Novoa R, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118.
2. Suk H, Lee S, Shen D, ADNI. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage*. 2014;101:569.
3. Schlegel T, Waldstein S, Vogl W, Schmidt-Erfurth U, Langs G. Predicting semantic descriptions from medical images with convolutional neural networks. Paper presented at: IPMI, 2015.
4. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proc. IEEE*. 1998;86(11):2278.
5. Krizhevsky A, Sutskever I, Hinton G. Imagenet classification with deep convolutional neural networks. Paper presented at: Advances in neural information processing systems (NIPS), 2012.
6. Prasoon A, Petersen K, Igel C, Lauze F, Dam E, Neilsen M. Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. Paper presented at: MICCAI, 2013.
7. Moeskops P, Viergever M, Mendrik A, de Vries L, Benders M, Isgum I. Automatic segmentation of MR brain images with a convolutional neural network. *IEEE Trans. Medical Imaging*. 2016;35(5):1252.
8. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. *Medical Physics*. 2017;44(2):547.
9. Setio A, Ciompi F, Litjens G, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE Trans. Medical Imaging*. 2016;35(5):1160.
10. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Paper presented at: MICCAI, 2015.
11. Peressutti D, Schipaanboord B, van Soest J, et al. How Effective Are Current Atlas Selection Methods for Atlas-Based Auto-Contouring in Radiotherapy Planning? *Medical Physics*. 2016;43(6):3738.

www.mirada-medical.com

Mirada Medical Ltd. Oxford Centre for Innovation, New Road, Oxford, OX1 1BY| +44 (0)1865 261410 | enquiries@mirada-medical.com.